

CLAIMS

1. A system for measuring the similarity between protein profile matrices in order to predict a protein three-dimensional structure, wherein
5 said profile matrix consists of a group of profile columns containing occurrence probabilities every amino acid type at a respective locations of amino acid residues in a multiple alignment in which amino acid sequences in a plurality of related proteins are aligned in multiple,

10 said system for measuring the similarity comprises the following means:

(a) means for preparing two profile matrices of a query profile matrix formed based on a plurality of proteins including proteins having three-dimensional structures to be predicted and a subject profile matrix formed based on a plurality of proteins having known three-dimensional structures;

15 (b) means for calculating correlation coefficients between the respective profile columns in said query profile matrix and the respective profile columns in said subject profile matrix with respect to full or partial combinations of both the respective profile columns; and

20 (c) means for forming a score matrix consisting of said correlation coefficients.

2. A system for predicting a protein three-dimensional structure characterized in using a score matrix formed through a system set forth in
25 claim 1.

3. A program for enabling a computer to function as a system for measuring the similarity between protein profile matrices in order to predict a protein three-dimensional structure, wherein

30 said profile matrix consists of a group of profile columns containing occurrence probabilities every amino acid type at a respective locations of amino acid residues in a multiple alignment in which amino acid sequences in a plurality of related proteins are aligned in multiple,

35 said system for measuring the similarity comprises the following means:

(a) means for preparing two profile matrices of a query profile matrix

formed based on a plurality of proteins including proteins having three-dimensional structures to be predicted and a subject profile matrix formed based on a plurality of proteins having known three-dimensional structures;

- 5 (b) means for calculating correlation coefficients between the respective profile columns in said query profile matrix and the respective profile columns in said subject profile matrix with respect to full or partial combinations of both the respective profile columns; and
- 10 (c) means for forming a score matrix consisting of said correlation coefficients.

4. A computer-readable recording medium storing a program set forth in claim 3.